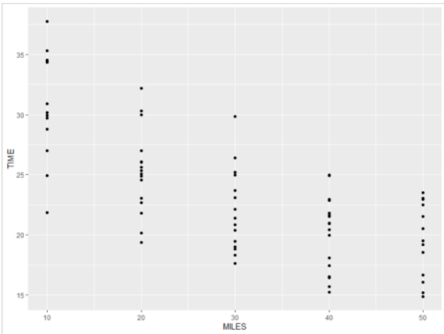
polynomial regression II

January 24, 2024

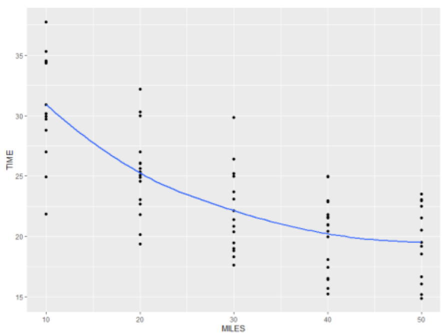
1

(hypothetical) data reminder



2

with a best-fitting curve



3

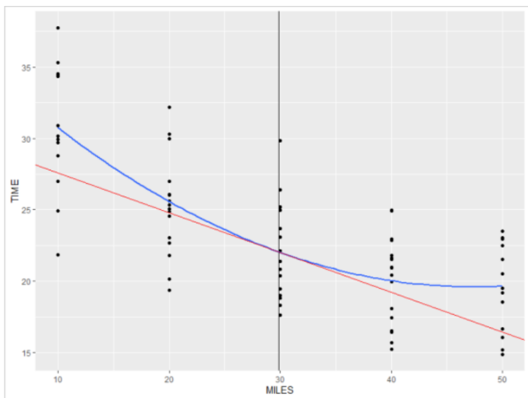
parameter estimates

- miles is mean-centered
- the how-to for this will be in this week's drill

```
             Estimate Std. Error t value
(Intercept) 22.053552   0.581916  37.898
MILES.c     -0.279100   0.027734 -10.063
M2           0.007941   0.002331   3.407
```

- the slope for MILES.c is the relationship between MILES and TIME only when MILES.c = 0 (i.e., at the mean) (*point slope*)
- the slope for M2 is (half) the rate at which the slope of MILES changes for each unit increase in miles

4



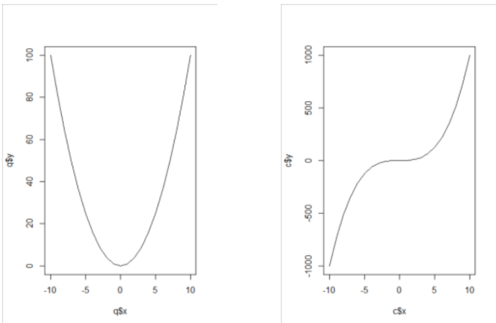
5

how to decide what power predictor to add

- if the relationship changes slope (e.g., from positive to negative) once, a squared predictor may work
- if the relationship changes slope twice (e.g., from + to – to + again), a cubed predictor may work

6

polynomials

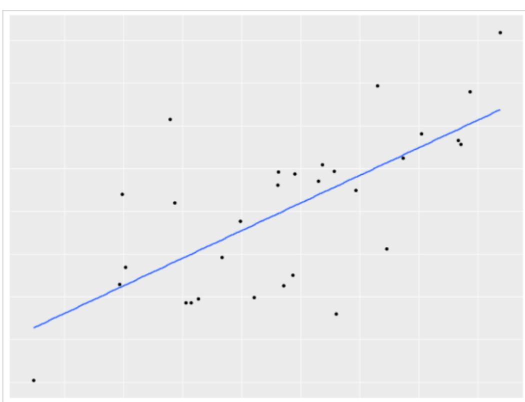


7

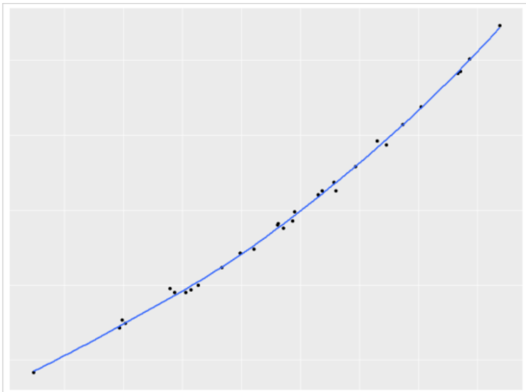
monotonic vs nonmonotonic relationships

- a *monotonic* relationship is one in which as one variable increases, so does the other (or vice versa)
- this may be linear or nonlinear
- a *non-monotonic* relationship is one in which the direction of the relationship changes as the value of the predictor changes

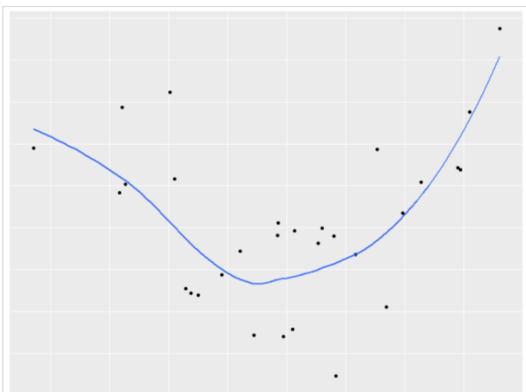
8



9



10



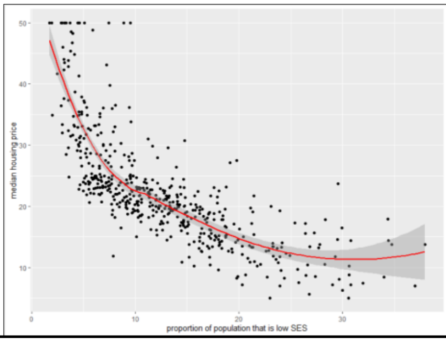
11

options with monotonic nonlinear relationships

- instead of adding a power predictor, you can simply transform either the predictor and/or the outcome to linearize the relationship
  - which transformation? see Tukey & Mosteller's bulging rule

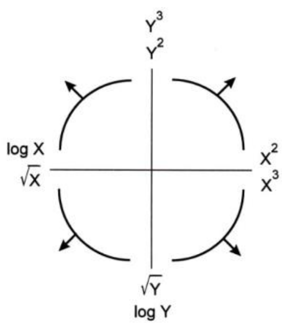
12

SES & housing prices



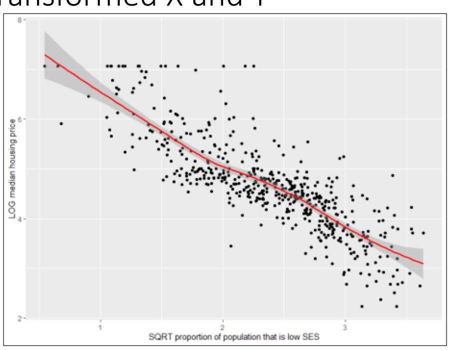
13

how to transform?



14

transformed X and Y



15

options with monotonic nonlinear relationships

- instead of adding a power predictor, you can simply transform either the predictor and/or the outcome to linearize the relationship
  - which transformation? see Tukey & Mosteller's bulging rule
- you can add a power predictor
  - pro: you can accommodate the changing X-Y relationship
  - con: just like with interactions, the interpretation of slopes is complex
- spline regression (read about it here if you're curious)

16

how do you know if you should add a power predictor?

- in an ideal world, a theoretical prediction will guide your modeling
- but you should look at your data
- scatterplots, esp. with the `geom_smooth()` function, will help you visualize what's going on
- as always, be clear in how you decided to analyze data; don't HARK (**h**ypothesize **a**fter **r**esults are **k**nown) – clearly identify exploratory analyses as such ... and then replicate!

17

significance testing polynomial slopes

- for the following model

$$Y = b_0 + b_1X + b_2X^2$$

- what's the null hypothesis for b_2?
- $b_2 = 0$ (precisely, $\beta_2 = 0$)

$$Y = b_0 + b_1X + 0X^2$$

18

interpretation caution

- the significance test for b_1 is a test of only the linear slope *at a particular value of X*; it is NOT a test of the main effect of X

19

a suggested protocol

- if you are interested in both the linear and quadratic effect of a predictor
- do things sequentially
- fit a one-predictor model first (and interpret b_1)

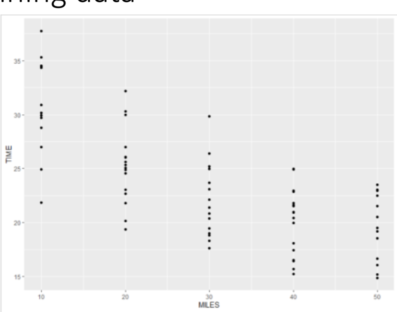
$$Y = b_0 + b_1X$$

- then add in the power predictor (and interpret b_2)

$$Y = b_0 + b_1X + b_2X^2$$

20

applying the protocol to the running data



21

step 1: the linear-only model

```
            Estimate       SE      t Pr(>|t|)
(Intercept) 23.55225  0.40609 57.998  < 2e-16
MILES.c     -0.27980  0.02956 -9.466 1.35e-14


SSE = 1029.0, Error df: 78, R-squared:  0.5346
```

In general, there is a strong negative linear relationship between miles of training per week and 5K times, b = -0.28, t(78) = -9.5, p < .001, R^2 = .53.

22

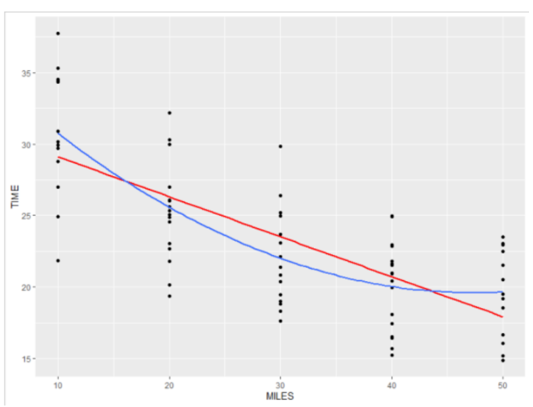
step 2: adding the quadratic term

```
             Estimate        SE       t Pr(>|t|)
(Intercept) 22.053552  0.581916  37.898  < 2e-16
MILES.c     -0.279100  0.027734 -10.063 1.09e-15
M2           0.007941  0.002331   3.407  0.00105

SSE = 894.2, Error df: 77, R-squared = 0.5956
```

The quadratic relationship is significant, b = 0.008, t(77) = 3.4, p = .001, ΔR^2 = .051. The relationship between training and 5K times is less negative as training increases.

23



24

examples in the literature (#1)

The Too-Much-Talent Effect: Team Interdependence Determines When More Talent Is Too Much or Not Enough

Psychological Science
2014, Vol. 25(8) 1581–1591

Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797614537280
pss.sagepub.com
SAGE

Roderick I. Swaab[1], Michael Schaerer[1], Eric M. Anicich[2], Richard Ronay[3], and Adam D. Galinsky[2]
[1]Organisational Behaviour Area, INSEAD, Fontainebleau, France; [2]Management Department, Columbia University; and [3]Department of Social and Organizational Psychology, VU University Amsterdam

25

examples in the literature (#1)

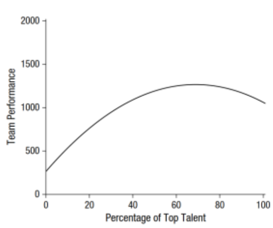
Table 2. The Impact of Talent on Football Teams' Performance in Study 2 (n = 415)

Predictor	Model 1	Model 2
Talent	1.84*** (0.16)	4.58*** (0.11)
Talent-squared	—	−4.26*** (0.49)
Roster size	0.04*** (0.01)	0.04*** (0.01)
Games played	0.03*** (0.01)	0.02** (0.01)
Intercept	4.63*** (0.11)	4.58*** (0.11)

Note: Standard errors are reported in parentheses. The corrected quasi-likelihood under the independence model criterion was 3,400.13 for Model 1 and 2,979.62 for Model 2.
**p < .01. ***p < .001.

26

examples in the literature (#1)



27

examples in the literature (#1)

Results were consistent with the lay intuition documented in Studies 1a and 1b, in that the linear relationship between talent and team performance was positive and significant (Table 2, Model 1). However, Study 2 also revealed a significant quadratic effect of top talent: Top talent benefited performance only up to a point, after which the marginal benefit of talent decreased and turned negative (Table 2, Model 2; Fig. 2). The linear and curvilinear effects remained significant when control variables were omitted ($b = 5.95$, $SE = 0.42$, $p < .001$, and $b = -4.98$, $SE = 0.57$, $p < .001$, respectively).

28

examples in the literature (#2)

Judgment and Decision Making, Vol. 11, No. 4, July 2016, pp. 352–360

Are neoliberals more susceptible to bullshit?

Joanna Sterling[*] John T. Jost[†] Gordon Pennycook[‡]

Abstract

We conducted additional analyses of Pennycook et al.'s (2015, Study 2) data to investigate the possibility that there would be ideological differences in "bullshit receptivity" that would be explained by individual differences in cognitive style and ability. As hypothesized, we observed that endorsement of neoliberal, free market ideology was significantly but modestly associated with bullshit receptivity. In addition, we observed a quadratic association, which indicated that ideological moderates were more susceptible to bullshit than ideological extremists. These relationships were explained, in part, by heuristic processing tendencies, faith in intuition, and lower verbal ability. Results are inconsistent with approaches suggesting that (a) there are no meaningful ideological differences in cognitive style or reasoning ability, (b) simplistic, certainty-oriented cognitive styles are generally associated with leftist (vs. rightist) economic preferences, or (c) simplistic, certainty-oriented cognitive styles are generally associated with extremist (vs. moderate) preferences. Theoretical and practical implications are briefly addressed.

Keywords: political ideology, neoliberalism, cognitive style, cognitive ability, bullshit receptivity

29

examples in the literature (#2)

Table 2: Linear models predicting bullshit receptivity.

	Model 1	*Model 2*	*Model 3*
Free Market Ideology	.006 (.003)*	.002 (.003)	.002 (.003)
Need for Cognition		.001 (.003)	
Heuristics and Biases		–.895 (.289)**	
Faith in Intuition		.009 (.003)**	
Numeracy			–.345 (.211)
Verbal Intelligence			–.837 (.389)*
Abstract Reasoning			–.418 (.326)

Note. We calculated bullshit receptivity as the average profundity rating of 30 statements that were rated on a scale from 1 (*Not at all profound*) to 5 (*Very profound*). "Heuristics and biases" was measured with the use of a battery of decision-making problems; higher scores indicate *less* reliance on heuristics, biases, and incorrect intuitions (Toplak et al., 2011). Need for cognition and faith in intuition were both measured in terms of self-report scales; higher scores indicate greater need for cognition and faith in intuition, respectively. Free market ideology was measured with the use of a five-item scale; higher numbers indicate stronger endorsement of free market ideology.

$^{\dagger} p < .06$, $^{*} p < .05$, $^{**} p < .01$.

30

examples in the literature (#2)

3.2.1 Quadratic effects of ideological extremity

To explore the possibility that ideological extremists would be more susceptible to bullshit than moderates, we centered free market ideology scores at the mean and computed a quadratic term. In an initial model, we observed a significant quadratic relationship such that those who were moderate in terms of their support for the free market appeared to be more susceptible to bullshit than extremists in either direction, $b = -.00027$, $SE = .00012$, $t(160) = -2.25$, $p = .026$ (see Figure 1 and Table 3).

31

examples in the literature (#2)

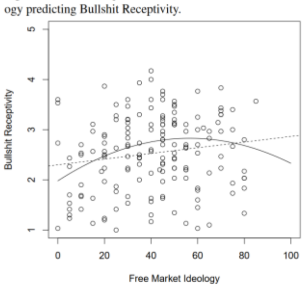
Table 3: Quadratic models predicting bullshit receptivity.

	Model 1		*Model 2*		*Model 3*	
Free Market Ideology						
Linear effect	.005	(.003)†	.002	(.003)	.003	(.003)
Quadratic effect	–.00027	(.00012)*	–.00016	(.00011)	–.00012	(.00012)
Need for Cognition			.001	(.003)		
Heuristics and Biases			–.807	(.295)**		
Faith in Intuition			.009	(.003)**		
Numeracy					–.305	(.215)
Verbal Intelligence					–.775	(.394)*
Abstract Reasoning					–.398	(.327)

32

examples in the literature (#2)

Figure 1: Linear and quadratic effects of Free Market Ideology predicting Bullshit Receptivity.



33